NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

A COMPARISON OF PASSWORD TECHNIQUES

by

Mark G. Beedenbender

March 1990

Thesis Co-Advisors: William J. Haga
Moshe Zviran

Approved for public release; distribution is unlimited

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Naval Postgraduate School
6c. ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5000	7b. ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. | PROJECT NO. | TASK NO. | WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)

A COMPARISON OF PASSWORD TEC NIQUES

12. PERSONAL AUTHOR(S)

Beedenbender, Mark G.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Master's Thesis	FROM ______ TO ______	1990, March	86

16. SUPPLEMENTARY NOTATION

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Information System Security; Authentication; User Identification; Passwords; Cognitive Passwords (KR)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A widely used access control mechanism is the password. Passwords are normally composed of a meaningful detail, such as a name of a person or a sequence of numbers such as a birthdate. Any person attempting to gain unauthorized access to a system might need only to look at a personnel record or associate with the person holding the desired password in order to discover the password. Therefore, there is a compromise between user memorability and security of a system. Exploration into other methods of user authentication and access control is desired to discover a better alternative to the traditional password system. The alternatives are system-generated passwords, pronounceable passwords, passphrases, cognitive passwords and authentication by word association. These methods are discussed and examined. The results from this study show that cognitive passwords and authentication by word association are superior to other methods in access control. Keywords; theses,

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Prof. William J. Haga	22b. TELEPHONE (Include Area Code) (408) 646-3094 | 22c. OFFICE SYMBOL Code AS/Hg

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

☆ U.S. Government Printing Office: 1986—606-24.

UNCLASSIFIED

Approved for public release; distribution is unlimited

A Comparison of Password Techniques

by

Mark G. Beedenbender
Lieutenant, United States Navy
B.A., University of Notre Dame, 1984

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN INFORMATION SYSTEMS

from the

NAVAL POSTGRADUATE SCHOOL
March 1990

Author: ______________________
Mark G. Beedenbender

Approved by: ______________________
William J. Haga, Thesis Co-Advisor

____
Moshe Zviran, Thesis Co-Advisor

____
David R. Whipple, Chairman
Department of Administrative Sciences

ABSTRACT

A widely used access control mechanism is the password. Passwords are normally composed of a meaningful detail, such as a name of a person or a sequence of numbers such as a birthdate. Any person attempting to gain unauthorized access to a system might need only to look at a personnel record or associate with the person holding the desired password in order to discover the password. Therefore, there is a compromise between user memorability and security of a system. Exploration into other methods of user authentication and access control is desired to discover a better alternative to the traditional password system. The alternatives are system-generated passwords, pronounceable passwords, passphrases, cognitive passwords and authentication by word association. These methods are discussed and examined. The results from this study show that cognitive passwords and authentication by word association are superior to other methods in access control.

TABLE OF CONTENTS

I. INTRODUCTION

A. COMPUTER SECURITY OVERVIEW

1. The Need for an Effective Authentication Method

The dependence of organizations upon computer systems necessitates the development of an authentication technology that provides security (Spender, 1987). Generally, the value of a system's access control is not given much thought until a major damage or loss occurs. Computer security policies are best established before, not after, an intrusion happens (Hoffer and Straub, 1989).

The growth of computer crime also necessitates improved authentication and security methods. In 1979, computer crime losses in corporations were estimated to be $100 million a year (Denning, 1979). By 1989, the figure had reached $500 million annually for 72 of the Fortune 500 companies (Hoffer and Straub, 1989). These numbers do not include all computer crimes that actually had taken place. Many went unreported for various reasons. Some are yet undiscovered. Companies feel that reporting the loss or damage would alert possible perpetrators to their inadequate security measures (Hoffer and Straub, 1989).

Computer crimes are not limited to pilfering company assets. There are many ways to commit criminal acts:

1. Damaging the computer center physically so that the hardware is no longer usable.

2. Using the computer system to modify, manipulate or delete computer software, resulting in monetary or strategic gain for the individual.

3. Using a computer to aid in the execution of a crime (e.g., programs designed to assist in infiltrating another system or other programs that produce forged statements to encourage investment in an individual's company).

4. Preying on people's ignorance to convince them to invest in a computer described as having programming or capabilities that do not exist. (Parker, 1984)

Also, computer crime is not always done for financial gain. Many "hackers" feel challenged to prove that no computer system is secure. While they may just play pranks once they have gained access, this access results in lost CPU and user time (Wood, 1983).

Finally, computer viruses introduced into computer systems and networks of computer systems can create many different problems. They can cause an operating system to "lock up" the computer system. They can attach themselves to programs and cause massive deletion of data. Viruses are insidious because they do not have to occur right away; they can lie dormant and be triggered later by such things as a system clock (Hoffer and Straub, 1989).

Since computers are becoming necessary to business, computer crime can only be expected to increase. Losses may be expected to increase because more costly data and information will continue to be entrusted to the computer. Finally, computers can be expected to be used more for illegal purposes--bookmaking, fraud and various other scams

(Parker, 1984). One of the best ways to counteract this barrage of crime will be to establish an increased awareness of computer security.

Computer security uses the technology, procedures, techniques and policies to guarantee the safety of not only the computer systems, but also the information stored within them. Security also involves limiting access to authorized users only (Ware, 1984). Each organization must evaluate what steps they need to take to make their system secure. Some computers are highly complex and need more elaborate security models to protect them whereas a home microcomputer will not need such intricacy to make it secure. Unless each organization establishes policies concerning computer security, it is difficult to determine if that system is secure for that company (Landwehr, 1981).

2. The Different Phases of Computer Security

There are many ways to approach computer security. First, external measures can be taken. Examples of such methods are making physical access to computer terminals difficult by use of guards, locks or some type of token (Ahituv et al., 1987). Second, access control is used to prevent unwarranted intrusions. Access control ensures that unauthorized people do not gain entry into a system, as well as preventing an authorized user from performing a function inside the system that is not allowed (Wood, 1983). Finally, there are internal controls to prevent illegal

tampering with data. These controls are designed to prevent users from accessing segments of memory to which they are not authorized. While access control is one method to provide internal security and control, there are other specific methods that can be used in conjunction with access controls to thwart the intended intruder (Denning and Denning, 1979).

This thesis focuses on an examination of various user authentication methods for improving computer security. It discusses and evaluates system generated passwords, pronounceable passwords, passphrases, cognitive passwords and authentication by word association as alternatives to the accepted traditional password system. A comparison of these authentication methods will show which method best balances the need to have a password system that is difficult for an unauthorized user to penetrate, yet is easy to use and remember for the legitimate user.

II. ROLE OF PASSWORDS AS AN AUTHENTICATION MECHANISM

A. PURPOSE OF PASSWORDS

A facet of computer security is ensuring proper user authentication. Password systems are the most commonly used authentication method. Those trusted with creating an effective password system view it as building a protective wall around an important physical asset. When a correct user identification and associated password are presented, the authorized user essentially passes over this logical fence to gain access to the computer system (Wood, 1987). As such, a password system is normally one of the first security measures used to deter unauthorized access to a computer system. Sometimes it may be the only method to stop illegal access. Therefore, it is important that this line of defense be as formidable as possible (Wood, 1983).

B. PASSWORD CHARACTERISTICS

A password is a combination of letters, numbers, special symbols or control characters that is used to verify that an authorized user is accessing the computer system (Wood, 1983). A password system could also be a longer string of such elements or a series of queries-and-responses in which each response was treated as an individual password. In order to make a system secure, there are several

characteristics that a password system should have. They are:

1. Ease of memorability. Passwords should not be written down.

2. Difficulty in guessing. This difficulty prevents discovery by an intruder. A password should not be associated with its user.

3. Ease of entry. They should not require a great deal of keyboard manipulation or expertise to enter the correct password.

4. Non-reusable. For instance, when a user changes his password, he should not be allowed to reuse a password he was assigned previously.

5. Non-susceptibility to spoofing. The password system should not be susceptible to a phony software program where the user is led to believe that he is logging in, when in reality a program simulating the login procedure is copying his password and user id. This form of spoofing is known as the Trojan horse.

6. Tested. A password system should be thoroughly evaluated and, once accepted, should be easy to implement.

7. Inexpensive. Password systems are generally an inexpensive way to provide protection, so the costs should not be prohibitive to install a given system. (Ahituv et al., 1987)

By far the greatest challenge in establishing an effective password system is to construct a system that has the first two characteristics--easy to remember, yet difficult to guess. Usually, some tradeoff is made concerning these two characteristics. On one erd of the spectrum, passwords can be made easy to remember. A user chooses a password that relates to him--a person, place or object. Such a choice makes it easy for someone else to guess the password within a few attempts (Barton and Barton, 1984).

However, as a password is made more complex, i.e., unrelated to the user, through the use of random characters for example, it is more difficult for the user to remember. As a result, compromise may occur because a user writes it down (Spender, 1987). Similarly, if an extended query-and-response routine is used, it may become tiring to the user, making it unpopular (Wood, 1983). It appears that with any password system some tradeoff must be made between these two characteristics.

C. TRADITIONAL PASSWORD SYSTEMS

In order to establish a metric on which to base alternative password systems, the traditional password system will be analyzed. Since this system is the most commonly used in existing operating systems it will serve as a benchmark on which to base alternative password methods. In this system a user is given a user identification (userid) and is instructed to select a password that will allow him access to the computer system. Normally a user is not restricted in any way in choosing this password; therefore, he generally picks a password because it is meaningful to him. In order to make this system secure, the password should be at least six characters in length. This should prevent an intruder from using "brute force" methods--trying all possible combinations, of say, four characters--to get access to the computer system (Wood, 1983). Also, studies have shown that people can readily

remember expressions up to only seven characters in length (Miller, 1956). For ease of memory and use, a user frequently will choose a password that is either a familiar name or a word found in the dictionary. This choice, of course, plays into the hand of an intruder as this reduces the possible combinations drastically. If, however, a password of more than six random characters is chosen, then it is more secure although not as acceptable to the user (Wood, 1983). An organization should establish a policy that strikes a balance between user-friendliness and susceptibility to compromise (Wood, 1983). After a password is chosen, the user will be required to use it every time he desires to log on to the system. Therefore, an important part of the security plan is that the password is changed after a period of time to avoid compromise. An intruder may discover a userid and simply try various guesses before hitting upon the matching password. Once again it is up to the organization to set policy as to how long a time period should be before the password is changed. In addition to changing the password, it is important that old passwords not be reused. For instance, a user must not use one password for a month, then switch to his other password, alternating between two passwords. If care is taken in selecting the password then the password system will be more effective at thwarting an unauthorized user (Wood, 1983).

An important element of this system, as with any other password system, is user education. People must be led to understand the importance of a password system and the steps to follow to ensure that the password chosen keeps the system secure. If the users do not believe in or understand the password system then it will become useless and ineffective (Wood, 1983).

Unfortunately, there are disadvantages to the traditional password system. They are:

1. If users decide to make the password as secure as possible--completely random--they tend to write it down so as not to forget it. By doing so they are leaving it open to compromise.

2. The user does not put effort into selecting a password, choosing a familiar name or trivial association, making it easy for an intruder to figure out.

3. Even if a "good" password is chosen, the user may key it in slowly or allow someone to watch as he keys it in, the password can be compromised.

4. This password system can be infiltrated by a sophisticated intruder through the operating system. Such an intruder can either find the password table and decipher it or use a spoofing technique to capture the password. (Ahituv et al., 1987)

The way to solve these problems is through better encryption techniques within the operating system and better user education. Is this the best method available? Or, is there a better alternative? Many such methods have been suggested and a description of several of these are provided in the next chapter.

III. ALTERNATIVE PASSWORD METHODS

While this chapter is not an exhaustive list of password system alternatives, the different alternatives presented are representative of various authentication methods. The first three methods presented are basically a modification of the traditional password system. Then a discussion follows on cognitive passwords and authentication by word association, new authentication approaches, not as closely linked to the traditional password system.

A. SYSTEM-GENERATED PASSWORDS

In the traditional password system, the user selected his own password, which was usually connected to his biography and therefore, guessable by outsiders. With a system-generated password, a system security administrator controls the selection of the password (Menkus, 1988). Within this method, it is common that a program creates passwords for users. Normally, a pseudo-random number generator arbitrarily creates a string of alphanumeric characters as the password (Menkus, 1988).

The system security administrator is responsible not only for selection of the passwords, but he must also ensure the proper distribution and use of the passwords. He is responsible for ensuring that passwords are changed frequently and that expired passwords are disposed of

properly and are not reused (Menkus, 1988). Depending upon the frequency of change, he may also need to change his generation program if it appears a pattern is developing among the passwords being created. In this method, the system security administrator is vital to ensuring for an effective authentication system (Menkus, 1988).

As discussed earlier, passwords should be chosen so they are difficult to guess or figure out. The advantage of the system-generated method is that it makes it more difficult for an intruder to penetrate it than is possible with the traditional password method (Panns and Herschberg, 1987). However, it will be more difficult for the user to remember since there is no meaningful relation to the user (Menkus, 1988). As a result, the high degree of complexity may cause the user to write down or even forget the password, thus failing to provide secure access control (Spender, 1987). Also, this method may result in friction between the user and the administrator's need to meet security requirements. A consequence may be that users will rebel in order to gain a system that is easier to use and remember (Panns and Herschberg, 1987).

B. PRONOUNCEABLE PASSWORDS

Pronounceable passwords consist of a string of alphanumeric characters that do not spell a word but rather, when pronounced or seen by a user, form a memorable string of characters. For instance, the word "operation" could be

made a pronounceable password by changing it to "oper8ion." Similarly, the famous Shakespearean quotation, "to be or not to be" changes to "2BORNOT2B" (Barton, 1984). These passwords can be either user-generated or system-generated. However, normally they are system-generated. As such, the administration of this password method would be the same as described in the system-generated password system.

The advantage of this method is that the password is not likely to be connected to the user's lifestyle. Also, since it is not as complex as a system-generated password, it should be easier to remember. However, because it is pronounceable does not mean it will be easy to remember, so some problems may still exist (Panns and Herschberg, 1987). Because users may be involved in the selection process, pronounceable passwords do not have to be system-generated. Also, there may not be as much user hostility toward this method as there is toward a system-generated authentication system.

C. PASSPHRASES

A variation of the traditional password system is the extended password or passphrase. Because it becomes more difficult to guess or find out a password as its length increases, the passphrase was designed to form a compromise between ease of memorability and difficulty in figuring out. The longer, extended password of 30 to 80 characters becomes difficult to guess (Porter, 1982). Unlike system-generated

passwords, the passphrase is generated by the user himself. This allows the user to choose a passphrase that is meaningful to him for ease of memory. In the passphrase method the sheer length of the passphrase provides the desired security, so having the passphrase unrelated to the user is not as stringent a requirement. The following example shows how length thwarts a possible intruder. If a user were to use a minimum of 30 alphabetic characters, over 1,000,000,000,000 possible combinations exist. This definitely makes the brute force attempt of trying all possible character combinations a formidable obstacle to an intruder (Pfleeger, 1989).

As long as the user avoided selecting a trivial passphrase, such as family names or the 26 letters of the alphabet, the ability to guess the passphrase would be unlikely (Porter, 1982). It seems that this method may be the one that finally resolves the conflict between the required characteristics of ease of memorability versus the difficulty in guessing. However, as stated earlier, a human being has difficulty remembering strings greater than seven characters (Menkus, 1988). Since the passphrase must be entered exactly, the question may be asked whether a human being can remember exactly a long string that has some meaning for him. Another problem may arise with the frequent user. Even though he may recall his passphrase without writing it down, he may become upset at the prospect

of typing 30 to 80 characters every time he desires to use the system (Porter, 1982).

D. COGNITIVE PASSWORDS

1. Description

Another alternative to the traditional password system is a method that lengthens the user identification process. Instead of a user entering just one password, he is required to enter several passwords, one at a time when prompted by the computer. One type of such a system is cognitive passwords. Cognitive passwords are passwords based on an individual user's perceptions, personal interests and personal history. These passwords are based on information that others would not commonly associate with the user, nor that could easily be found in personal records (Haga et al., 1989).

2. Implementation

A cognitive password system combines both system-generated and user-generated characteristics. It is system-generated in that the security administrator creates questions that would be used to stimulate a response from a user. The exact responses to these questions would entirely be user-generated. As such, the password system is set up basically as an access quiz. If the user responds correctly to a series of questions concerning himself, then he would be authorized access to the system (Haga et al., 1989).

In order to make the security system effective, the administrator needs to choose non-trivial questions as the stimulus for user responses. If trivial questions such as "What is your name?" are chosen, then an intruder will more easily break into the system than if "What is your favorite vacation place?" is used, for example.

Like the other password methods described, the responses or cognitive passwords would need to be entered exactly for a user to gain access. Because the responses to the questions necessarily vary in length, cognitive passwords would have, however, no preset length associated with them. They also would be regular words as opposed to a random string of alphanumeric characters.

3. Advantages

Since the cognitive password is significant to the user, but not readily associated with him, it is easy for him to remember, but difficult for an intruder to guess or find out. The cognitive password may be of such length that a brute force method of trying all character combinations would be thwarted. Also, a cognitive password system requires several questions to be answered correctly, so this layering adds an additional degree of security.

4. Disadvantages

In the traditional password system it is difficult for a user to remember one password, therefore remembering many cognitive passwords would seem to be harder for the

user (Smith, 1987). Also, it is unlikely that a user would remember all of his responses so establishing an acceptable miss percentage may be difficult to do. If set too low, intruders may penetrate the system; if set too high, authorized users may be denied access.

5. Summary

A cognitive password system, if implemented correctly, seems to be less vulnerable than the traditional password system. It also provides for user-friendliness. Even though it may be more complicated to set up initially, the benefits of a better authentication system make it a viable alternative.

E. WORD ASSOCIATION

1. Description

Another method that requires a series of passwords to verify user identity is authentication by word association. In this alternative, the user constructs a list of cues and responses that would be unique to the individual. A trivial example would be the cue word "high" which would require the response "low" (Smith, 1987). Smith (1987) designed this model with the thought that an initial list of 20 cues and responses per user would be sufficient to allow flexibility in changing the cues presented to the user when logging in to the system. Depending upon the security of the system, a user would be required to give from one to several correct responses (Smith, 1987).

2. Implementation

The actual structure desired would be a single-word cue and a one-word response. Doing so allows for ease of memory for the user. Similarly, the user is responsible for constructing all 20 cues and responses making it user-friendly (Smith, 1987).

In order to make this method a stronger impediment to intrusion, the word associations should be as non-trivial as possible. A list of 20 opposites would be easy to penetrate (Smith, 1987). To make construction of the list easier and to make it easier for the user to remember the responses, it is helpful for the user to choose one central theme (Smith, 1987). For example, United States presidents may be chosen as the theme. Cues may include cherry and honest and have responses of Washington and Lincoln, respectively.

Finally, while the user is expected to generate the correct response to gain access to the system, there is no requirement that he remember the cues or the central theme, if any, for the word association method to be successful (Smith, 1987).

3. Advantages

Smith postulated that there were several advantages to this method:

1. The responses would be easy to remember.
2. Without knowledge of the theme and non-trivial associations, the responses would be resistant to intrusion.
3. Since the cues and responses are selected by the user, there would be little user resistance to such a method.
4. The cues and responses would uniquely identify each individual user.
5. If a need arose to change a cue and response, it could easily be altered without altering or compromising the rest of the list. (Smith, 1987)

4. Disadvantages

If a user is not careful in constructing his word associations, the responses may be easily guessed. Also, the user may be tempted to write down the cues and responses or the central theme since there would be so many responses to remember. This would lead to compromise. And, like cognitive passwords, a user would likely not remember all the responses so an acceptable margin of incorrect responses would have to be established.

5. Summary

Smith's word association model offers an alternative to the traditional password system as the word association model is more robust and would require a great deal of effort to penetrate. It is user-friendly in both design and use. Unfortunately, in Smith's study he only tested four users, so further testing may be required.

The next three chapters contain the methodology, findings and conclusions of examining the different password

methods and in determining the possible superiority of any one method over the others.

IV. RESEARCH METHODOLOGY

A. BACKGROUND

The purpose of this research is to determine if any of the six password methods is superior in ease of use, memorability and resistance to intrusion. The method used to conduct this research was by questionnaire. Several different questionnaires were used with the intent either to verify information from previous studies or to justify conclusions about new areas of study, such as cognitive passwords and authentication by word association.

B. METHODOLOGY

1. Instrumentation

To test the ease of recall for all methods, three forms of similar self-administered questionnaires were developed. A copy of each questionnaire is included in the Appendix. Two versions of the first questionnaire, Q1, were used (the differences are described in Section B.1.b of this chapter). Each respondent (user) answered one of the two versions of the first questionnaire and the third form of the questionnaire, Q1 and Q3. A significant-other (spouse, close friend or family member) completed the second form of the questionnaire, Q2.

a. Demographic Items

Both the Q1 and Q3 forms asked for four categories of responses. The first part of Q1 asked for the respondent's sex, years of computer usage, types of computers with which they were experienced (microcomputer, microcomputer linked to a mainframe and/or a mainframe terminal) and a respondent identifier--either their last four digits of their United States Social Security number or their Student Mail Center (SMC) box number. The Q3 form asked only for the Social Security number digits or the SMC box number, so that it could be matched with its Q1 counterpart. The Social Security digits and the SMC number were used to protect the identity of individual respondents in this study, yet were used to allow matching of the Q1, Q2 and Q3 forms. The SMC number allowed the respondent to be contacted somewhat anonymously (as he was addressed by SMC number and not by name) for the return of either a Q2 form or missing information from the Q1 or Q3 form.

b. Creation and Assignment of Passwords and Passphrases

The second part of Q1, but not Q3, asked each respondent to construct a password consisting of any combination of up to eight alphanumeric characters. The test group was urged to memorize and safeguard this password as they would any other password. They were then asked how they devised this password. Four choices were given: (1) the password represents a meaningful detail such as a name,

a date or a number; (2) the password represents a combination of such meaningful details; (3) the password represents a random choice of characters; or (4) some other means. The second part of Q1 contained a unique eight character password that was assigned to each respondent. This was the only difference in the versions of questionnaire, Q1. Fifty-five of the Q1 forms had a system-generated random alphanumeric password. Forty-eight of the Q1 forms had a system-generated pronounceable password. To distinguish between the two versions of Q1, the random alphanumeric form was designated Q1R and the pronounceable password form was designated Q1P. The respondents were urged to safeguard this password as well.

Also included in the second part of Q1 was a segment requesting each respondent to create a passphrase consisting of any combination of up to 80 alphanumeric characters. There was no requirement as to the minimum number of characters in the passphrase. The respondents were again urged to memorize and safeguard this passphrase as they would any other password. Then they were asked how they devised this passphrase. Five choices were given: (1) nonsensical phrase; (2) a quotation; (3) a piece of advice; (4) a common phrase; or (5) other means.

c. Cognitive passwords

The Q1 and Q3 forms are identical in their third part. In this section, 20 open response questions ask for

items of information that were described as cognitive passwords. These items fall into two categories of responses. The first group are six items of personal facts assumed to be known only by the respondent or someone socially close to the respondent, for example, elementary school attended, mother's maiden name or father's occupation. The second group is 14 opinion-based items which ask the respondent to choose a favorite item, for example, favorite vacation place, favorite restaurant or favorite fruit. Once again, it was assumed that these responses would be known only by the respondent or someone socially close to him.

d. Word Association

The final part of the Q1 form requested the user to come up with a list of 20 word associations. In formulating these 20 cues and responses, the respondents were not required to use a central theme throughout. There was no limitation or minimum number of alphanumeric characters in either the cues or responses. They were then asked to copy just the cues onto Q2 in the appropriate spaces to see if a socially close person would be able to figure out the responses.

e. Items for Recall of Passwords

On Q1 respondents were asked to create a password; on the second part of Q3 the same respondents were asked to recall this password. Q3 was administered to these

same respondents approximately three months after the administration of Q1. After asking each person to recall the password of his own making, each respondent was asked what method was used to recall his password. The following choices were given: (1) Writing it down, even though they were asked not to; (2) Memory recall; (3) The only password the respondent ever uses; and (4) Other means.

Next each respondent was asked to recall the assigned password given on the Q1 form. The respondents were again asked how they recalled the password: (1) Memory recall; (2) Writing it down--even if told not to; (3) If the password were pronounceable, had that aided in recalling the password; or (4) Other means.

Finally, the respondents were asked to recall their passphrase from their Q1 form. They were then asked whether they had: (1) Written it down; (2) Recalled it from memory; (3) Chosen a phrase that they use over and over again so it was easy to remember; or (4) Other means. Expectations were that of the four methods (user-generated passwords, system-generated random passwords, system-generated pronounceable passwords and passphrases) pronounceable passwords would be recalled the most often.

f. Items for Recall of Cognitive Passwords

In the identical Q3 version of the cognitive password section, the same respondents were asked the same questions again. As with the previous part, Q3 was

administered approximately three months after Q1. In examining a system of passwords based upon cognitive information, the correlation between the Q1 and Q3 cognitive responses is of interest. Expectations were high that there would be a high correlation, especially among the fact-based cognitive items.

g. Items for Recall of Word Associations

In the identical Q3 version of the word association section, the same respondents were asked to regenerate their list of 20 cues and responses. As with the other segments of Q3, this part was administered approximately three months after Q1. Expectations were that the respondents would be able to come up with few, if any, of their original cues and responses. As soon as the respondent had generated as many associations from memory as possible, they were given a list of their original 20 cues. They were then asked to write down as many of their responses as they remembered. If, at this point, they were still unable to remember their responses, they were given the central theme, if any, to aid them in correctly remembering their responses. Expectations were that the respondents would not remember many of their original cues and responses. However, once the cues were given to the respondents, it was expected that most would remember their responses and very few would need to be told their theme, if any, to help in figuring out their responses.

h. Items Concerning the Various Password Methods

The final section of Q3 requested the respondents to rank the various password methods--user-generated passwords, system-generated passwords, passphrases, cognitive passwords and authentication by word association--by ease of memory. The respondents were then asked to rank the methods as to how they liked them. Expectations were that the rankings would be similar. Also, the liking of methods with user involvement was expected to be higher than those that were system-generated.

i. Items to Tap Socially Close Knowledge

The Q2 significant-other form asked for two items of identifying data. It asked for the last four digits of the user respondent's Social Security number or the SMC number to be used for matching purposes. It then asked for the relationship of the Q2 significant-other respondent to the Q1 user respondent. The second part of the Q2 form repeated the 20 cognitive password questions that were in the third part of the Q1 form. The significant-other was asked to indicate what they thought the Q1 respondent would answer to each of the questions. They were then asked to complete the Q2 form without the help of the Q1 respondent. Of interest was the level of accuracy at which the Q2 significant-others could match the responses of the Q1 respondents. The assumption was that if someone socially close to a respondent had deficient

knowledge of personal cognitive data, then the likelihood of guessing by someone socially distant from the same respondent would be even less likely.

The third part of the Q2 form asked the significant-other to determine the responses to the cues written down by the respondents from the word association portion of their Q1. After attempting to figure out the responses without aid from the user respondent, the significant-other respondent was given a second chance. This time the respondent would inform their significant-other if there was a central theme to the associations and if so, what it was. Once again, of interest was the level of accuracy at which the Q2 significant-others could match the Q1 respondents. As with cognitive passwords, it was assumed that if someone socially close to the respondent was unable to figure out the responses, then the chances of an intruder figuring out the responses would be slim.

2. <u>Sample and Data Collection Design</u>

a. Q1 Response by the Respondents

The Q1 questionnaire was administered to 103 graduate students, the majority of which were majoring in management information systems. Of the respondents, 85% were male and 15% were female. Their level of experience with computers averaged five years. Twelve percent said they had no computer experience before starting graduate studies. Forty-eight percent reported that they used some

combination of microcomputer and mainframe, 32% said their computer experience was limited to microcomputers, while 8% claimed to have only used a mainframe.

b. Q2 Response by Significant-other

After completing the Q1 forms, the respondents were given the Q2 form. They were asked to write their SSN or SMC identifier on the Q2 form and to give the form to a significant-other of their choosing. They were then asked to return the Q2 forms after being completed by their significant-other. Q2 forms were returned by 85% of the respondents. Of the significant-others responding, 76% were spouses, 21% were friends and 3% were family members.

c. Q3 Response by Respondents

The Q3 version of the questionnaire was administered to the same user respondents approximately three months after the Q1 administration. Of the original 103 Q1 respondents, 100% participated in the Q3 administration.

C. TABULATION

Upon completion of the administration of the Q1, Q2 and Q3 questionnaires, the data were tabulated and analyzed using standard statistical methods. The mainframe software package used was SPSS-X, release 3.1. In Chapter V the findings and results from the questionnaires are summarized.

V. RESEARCH FINDINGS

A. FINDINGS

1. Recall of Self-generated Passwords

Of the 103 respondents, 27.2% were able to recall correctly the password they had created themselves three months earlier. Among the respondents who recalled their password, 42.9% said they remembered it without aid, 7.1% said they wrote it down even though they were instructed not to. 17.9% said it was the only password they ever used so it was easy to remember. Finally, 32.1% gave "other means" as the basis for recall. Most of them used some type of memory aid to help in recalling the password.

Table 1 shows how the respondents constructed their self-generated password. The majority, 67%, used some form of meaningful detail in creating their password. Figure 1 shows how many characters were used in construction of the self-generated passwords. There were eight spaces on the Q1 form and most, 54.4%, tended to use all the spaces when making up their passwords.

Table 2 shows the composition of the self-generated passwords. As expected, the respondents mainly used alphabetics in creating their passwords. More interesting is the fact that of the 28 respondents who recalled their password, 92.9% used alphabetics only. Furthermore, Table 3

TABLE 1

METHODS FOR SELECTING SELF-GENERATED PASSWORDS

METHOD	NUMBER	PERCENTAGE
MEANINGFUL DETAIL	46	44.7
COMBINATION OF MEANINGFUL DETAILS	23	22.3
RANDOM CHARACTERS	1	1.0
OTHER	33	32.0

TABLE 2

PASSWORD COMPOSITION

COMPOSITION	NUMBER	PERCENTAGE
ALPHABETIC ONLY	76	73.3
ALPHANUMERIC	24	23.3
ASCII CHARACTERS INCLUDED IN PASSWORD	3	2.9

shows that generally as the length increases, the ability to recall the password decreases. For instance, four of the eight four-letter passwords were recalled, thus half or a 50% recall occurred for four-letter passwords.

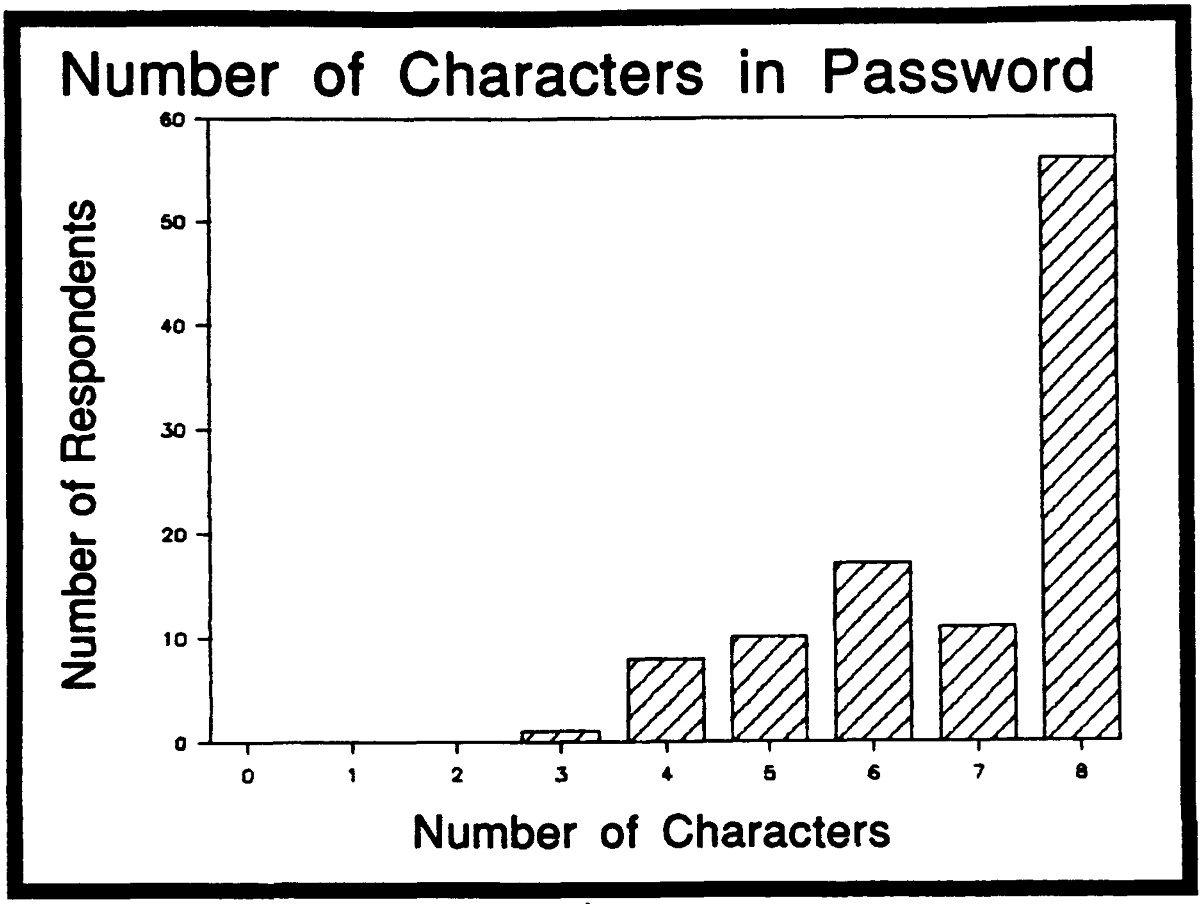

Figure 1. Number of Characters in Password

2. Recall of System-generated Random and Pronounceable Passwords

Table 4 reflects the ability of the respondents to recall either the assigned random alphanumeric password or the pronounceable password. As expected, fewer respondents were able to recall their system-generated password than their own self-generated password. However, the difference (28 versus 25 recalled) was not as large as expected. A

TABLE 3

PASSWORD LENGTH VS. MEMORABILITY

NUMBER OF CHARACTERS IN RECALLED PASSWORD	PERCENTAGE OF ALL PASSWORDS OF THAT LENGTH
3	0.0
4	50.0
5	20.0
6	41.0
7	27.0
8	21.0

TABLE 4

SYSTEM-GENERATED PASSWORD RECALL

TYPE OF PASSWORD	NUMBER ASSIGNED	NUMBER RECALLED	PERCENTAGE
PRONOUNCEABLE	48	18	37.5
RANDOM ALPHANUMERIC	55	7	12.7
TOTAL SYSTEM-GENERATED	103	25	

possible explanation for the closeness between the password methods is the recall of pronounceable passwords. Of the recalled system-generated passwords 72% were pronounceable.

The increased memorability of pronounceable passwords is further supported in Table 5. Sixty-seven percent of the respondents who recalled their pronounceable password stated that they remembered it simply because the assigned password was pronounceable. No one was able to recall their random password from memory. 85.7% had to write it down even though they had been instructed not to do so.

TABLE 5

METHOD OF RECALL FOR SYSTEM-GENERATED PASSWORDS

METHOD OF RECALL FOR PRONOUNCEABLE PASSWORD	NUMBER RECALLED	PERCENTAGE
BECAUSE IT WAS PRONOUNCEABLE	12	66.7
UNAIDED MEMORY	3	16.7
WROTE IT DOWN	3	16.7

METHOD OF RECALL FOR RANDOM PASSWORD	NUMBER RECALLED	PERCENTAGE
UNAIDED MEMORY	0	0.0
WROTE IT DOWN	6	85.7
OTHER	1	14.3

3. <u>Recall of Passphrases</u>

Of the 103 respondents, only 21.4% were able to recall the passphrase which they had created approximately three months earlier. As expected, a longer string of characters, even though it formed an expression familiar to a respondent, made it difficult to remember. Table 6 shows that the length of recalled passphrases was less than the length of all the passphrases originally created, but not dramatically so.

TABLE 6

PASSPHRASE LENGTH

	IN ALL PASSPHRASES	IN RECALLED PASSPHRASES
AVERAGE CHARACTER LENGTH	22.7	21.3
AVERAGE NUMBER OF WORDS IN PASSPHRASE	4.9	4.4

For those respondents who did remember their passphrase, Table 7 shows the method they used to recall it.

Table 8 shows the different methods used to construct the passphrases. No method was clearly preferred in creating the passphrase. As a matter of fact, each method, except for selecting a piece of advice, uniformly received little more than 20% usage by the respondents.

TABLE 7

METHOD OF PASSPHRASE RECALL

METHOD	NUMBER	PERCENTAGE
WROTE IT DOWN	2	9.1
UNAIDED MEMORY	9	40.9
COMMON PHRASE USED OFTEN BY RESPONDENT	5	22.7
OTHER	6	27.3

TABLE 8

METHOD OF CREATING THE PASSPHRASE

METHOD	NUMBER	PERCENTAGE
NONSENSICAL PHRASE	24	23.3
A QUOTATION	21	20.4
A PIECE OF ADVICE	10	9.7
A COMMON PHRASE	21	20.4
OTHER	27	26.2

4. Recall of Cognitive Passwords

a. Recall of Cognitive Passwords by Respondents

The overall average number of correct matches by the respondents on all cognitive passwords between Q1 and Q3 was 14.8 out of 20 correct responses or 74%. Figure 2 reflects this distribution. Of interest is the grouping of the respondents at the high end of the spectrum. While there are a few outliers at the low end of the spectrum,

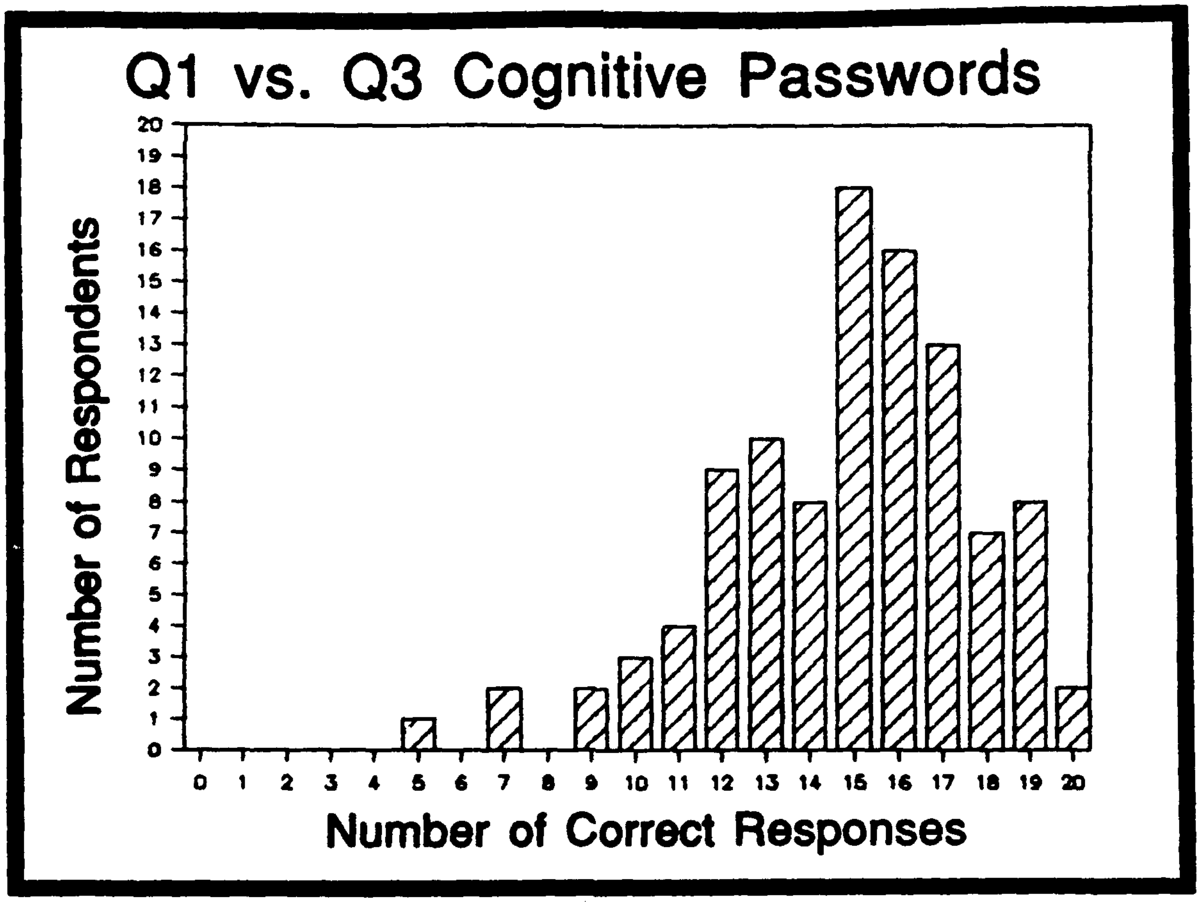

Figure 2. Q1 vs. Q3 Cognitive Passwords

which resulted in the somewhat low mean, 62.1% of the respondents had 15 or more correct responses. Of interest is the comparison of the level of these responses with the responses for the previously analyzed password methods. The best password response was 37.5% for the system-generated pronounceable passwords. On the cognitive password continuum, the number of correct matches for self-generated passwords is equivalent to getting 7.5 correct responses. Only three respondents scored that poorly on cognitive passwords.

Besides the overall high success rate, the respondent's performance for each individual question is of interest. As discussed in the research methodology chapter, the cognitive questions were split into six fact-based questions and 14 opinion-based questions. The success of the respondents in recalling cognitive passwords over a three month period is expressed in the percentage of correct matches that were produced on the Q3 form. Table 9 shows that the recall for the fact-based questions was high, 83.7%. Even the lowest cognitive question had a recall rate of 74.8%, twice the recall rate for any of the previous password methods.

The success rate for the recall of the opinion-based questions is lower than for the fact-based questions. The average percentage of correct responses was 70%. There was a fairly wide variance with the number of correct

TABLE 9

RESPONDENT MATCHING ON FACT-BASED COGNITIVE PASSWORDS

ITEM	NUMBER WHO MATCHED CORRECTLY	PERCENTAGE WHO MATCHED CORRECTLY
WHAT IS THE NAME OF THE ELEMENTARY SCHOOL FROM WHICH YOU GRADUATED?	87	84.5
WHAT IS THE NAME OF YOUR FAVORITE UNCLE?	89	86.4
WHAT IS THE NAME OF YOUR BEST FRIEND FROM HIGH SCHOOL?	87	84.5
WHAT IS YOUR MOTHER'S MAIDEN NAME?	96	93.2
WHAT WAS THE FIRST NAME OF YOUR FIRST BOYFRIEND/GIRLFRIEND?	77	74.8
WHAT IS THE OCCUPATION OF YOUR FATHER?	81	78.6

responses ranging from 49.5% to 87.4%. The questions that had the lowest success rate dealt with an individual's favorite restaurant, actor or actress, and choice of alternative profession. Two possible explanations for missing these questions are: (1) At the time of administration of Q1, the respondent may have wavered between a couple of answers, failing to remember which one he had chosen three months earlier and selecting a different answer on Q3; and (2) these questions call for answers that may have changed for the respondent since the administration of Q1. Therefore, the respondent may have answered the question according to his opinion at the time of the administration of Q3, as opposed to responding as he did when he first answered the question. Tables 10 and 11 show the results of the opinion-based cognitive questions.

b. Matching of Cognitive Passwords by Significant-Others

The average number of correct matches by significant-others on all cognitive passwords from the Q2 form was 7.6 out of 20 (38%). Figure 3 shows the distribution of the correct matches. The distribution approaches that of a normal curve. The distribution curve emphasizes the success rate of the significant-others and is skewed toward the lower end of the spectrum. However, it was not expected that as many significant-others would do as well as shown. An explanation for the success rate of the 17 significant-others who scored better than ten correct

TABLE 10

RESPONDENT MATCHING ON OPINION-BASED COGNITIVE PASSWORDS

ITEM	NUMBER WHO MATCHED CORRECTLY	PERCENTAGE WHO MATCHED CORRECTLY
WHAT IS THE NAME OF YOUR FAVORITE CLASS IN HIGH SCHOOL?	80	77.7
WHAT IS THE NAME OF YOUR FAVORITE MUSIC PERFORMER OR GROUP?	82	79.6
WHAT IS YOUR FAVORITE TYPE OF MUSIC?	89	86.4
WHAT IS THE NAME OF YOUR FAVORITE VACATION PLACE?	68	66.0
IF YOU COULD TRAVEL TO ANY COUNTRY IN THE WORLD, WHICH WOULD IT BE?	74	71.0
WHAT IS THE LAST NAME OF YOUR FAVORITE ACTOR OR ACTRESS?	60	58.3

TABLE 11

RESPONDENT MATCHING ON OPINION-BASED COGNITIVE PASSWORDS

ITEM	NUMBER WHO MATCHED CORRECTLY	PERCENTAGE WHO MATCHED CORRECTLY
WHAT IS YOUR FAVORITE FLOWER?	90	87.4
WHAT IS YOUR FAVORITE DESSERT?	68	66.0
WHAT IS YOUR FAVORITE VEGETABLE?	77	74.8
WHAT IS YOUR FAVORITE FRUIT?	68	66.0
WHAT IS YOUR FAVORITE COLOR?	77	74.8
IF YOU COULD CHANGE OCCUPATIONS, WHICH NEW OCCUPATION WOULD YOU CHOOSE?	56	54.4
WHAT IS THE NAME OF YOUR FAVORITE RESTAURANT?	51	49.5
WHAT IS THE LAST NAME OF YOUR FAVORITE COLLEGE INSTRUCTOR?	69	67.0

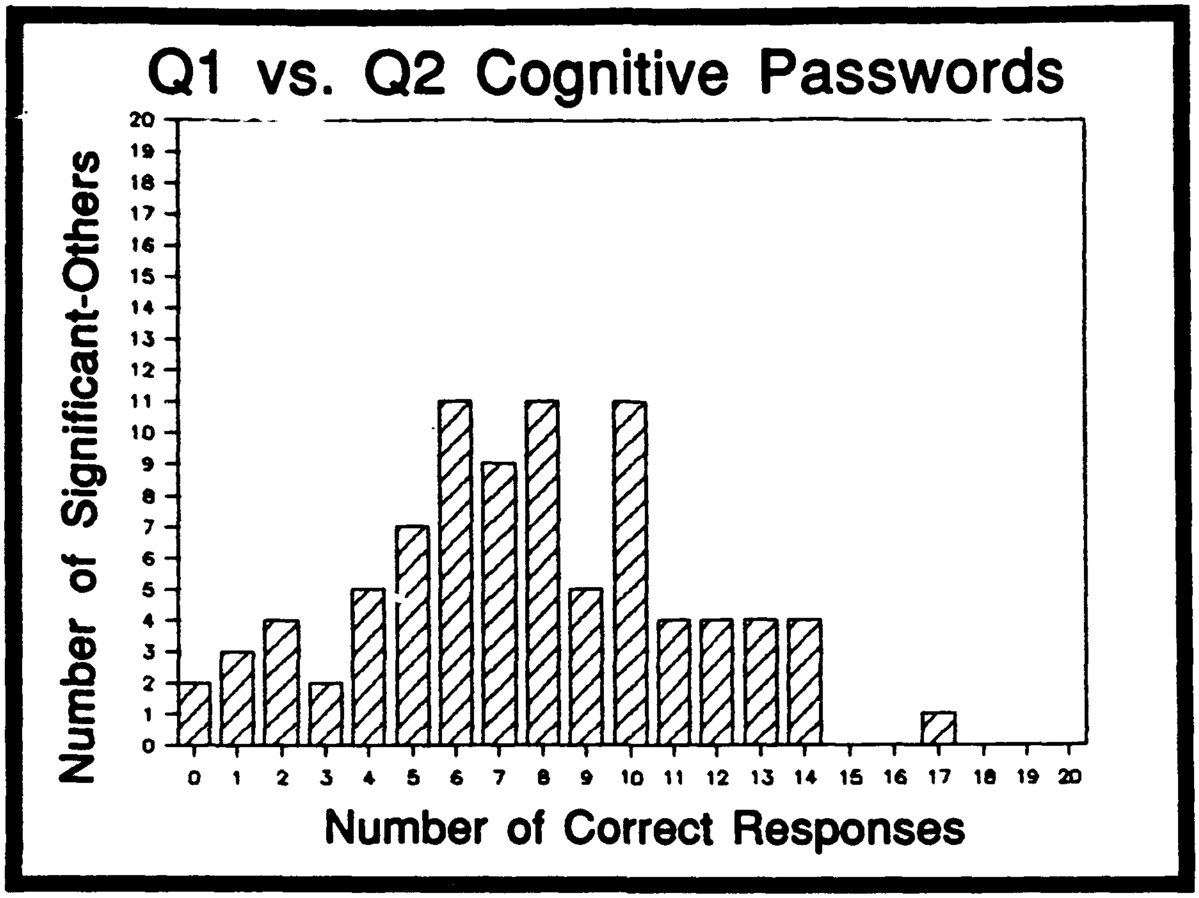

Figure 3. Q1 vs. Q2 Cognitive Passwords

answers may be that 13 of the 17 (76.5%, including the one scoring 17 correct matches) answered the same questions six months earlier when Hulsey was conducting similar research with some of the same respondents. This, also, may have affected the other results, too, as 60% of the respondents' significant-others were the same surveyed by Hulsey. In Hulsey's study, only one significant-other scored greater

than ten correct matches. Therefore, only four significant-others in this survey can be interpreted as scoring better than ten correct matches without help from the respondents. Interestingly enough the scores of these four significant-others were 11, 12, 13 and 14.

Significant-others in this study are assumed to be people who are close to the user respondents--spouses, close friends or family members. Yet, even they do not have correct knowledge, on an average of more than 40%, of the items on personal information and personal preferences of the respondents.

The difficulty the significant-others had in matching the cognitive passwords is confirmed in the average percentage score, 44.6%, for fact-based questions (Table 12). The assumption was made that the fact-based questions would be better known than would the opinion-based questions by significant-others. Nonetheless, even though the significant-others are precisely the people who should know better than anyone else the personal facts about the respondents, they, on average, knew less than half of the correct responses.

As expected, the significant-others knew less about the personal preferences of the respondents (Tables 13 and 14) then they knew about the respondent's personal facts. The average score of matches for the 14 opinion-based items is 4.6 (32.5%).

TABLE 12

SIGNIFICANT-OTHER MATCHING ON
FACT-BASED COGNITIVE PASSWORDS

ITEM	NUMBER WHO MATCHED CORRECTLY	PERCENTAGE WHO MATCHED CORRECTLY
WHAT IS THE NAME OF THE ELEMENTARY SCHOOL FROM WHICH YOU GRADUATED?	22	25.3
WHAT IS THE NAME OF YOUR FAVORITE UNCLE?	47	54.0
WHAT IS THE NAME OF YOUR BEST FRIEND FROM HIGH SCHOOL?	40	46.0
WHAT IS YOUR MOTHER'S MAIDEN NAME?	59	67.8
WHAT WAS THE FIRST NAME OF YOUR FIRST BOYFRIEND/GIRLFRIEND?	18	20.7
WHAT IS THE OCCUPATION OF YOUR FATHER?	47	54.0

TABLE 13

SIGNIFICANT-OTHER MATCHING ON
OPINION-BASED COGNITIVE PASSWORDS

ITEM	NUMBER WHO MATCHED CORRECTLY	PERCENTAGE WHO MATCHED CORRECTLY
WHAT IS THE NAME OF YOUR FAVORITE CLASS IN HIGH SCHOOL?	22	25.3
WHAT IS THE NAME OF YOUR FAVORITE MUSIC PERFORMER OR GROUP?	37	42.5
WHAT IS YOUR FAVORITE TYPE OF MUSIC?	44	50.6
WHAT IS THE NAME OF YOUR FAVORITE VACATION PLACE?	22	25.3
IF YOU COULD TRAVEL TO ANY COUNTRY IN THE WORLD, WHICH WOULD IT BE?	27	31.0
WHAT IS THE LAST NAME OF YOUR FAVORITE ACTOR OR ACTRESS?	25	28.7

TABLE 14

SIGNIFICANT-OTHER MATCHING ON
OPINION-BASED COGNITIVE PASSWORDS

ITEM	NUMBER WHO MATCHED CORRECTLY	PERCENTAGE WHO MATCHED CORRECTLY
WHAT IS YOUR FAVORITE FLOWER?	46	52.9
WHAT IS YOUR FAVORITE DESSERT?	34	39.1
WHAT IS YOUR FAVORITE VEGETABLE?	32	36.8
WHAT IS YOUR FAVORITE FRUIT?	32	36.8
WHAT IS YOUR FAVORITE COLOR?	51	58.6
IF YOU COULD CHANGE OCCUPATIONS, WHICH NEW OCCUPATION WOULD YOU CHOOSE?	25	28.7
WHAT IS THE NAME OF YOUR FAVORITE RESTAURANT?	20	23.0
WHAT IS THE LAST NAME OF YOUR FAVORITE COLLEGE INSTRUCTOR?	11	12.6

An assumption was made that the significant-others are the people in the best position to possess knowledge about the respondents. Of interest then is the ability to judge how much personal knowledge is held by socially close people. A further assumption is that the accuracy of personal knowledge would decrease as the social distance was increased.

To examine this social distance phenomenon, the average number of correct matches was calculated on the overall set of 20 cognitive questions for the three family members, the 66 spouses and the 18 friends. Unfortunately, there was not a larger sample of family members, so there may be some bias in the results. The average number of correct matches for family members was 12 (60%); for spouses it was 8.2 (41%); and for friends it was 4.7 (23.5%). The difference between each group is significant. However, it was no surprise that family members did the best as they have been exposed to the respondent for most of the respondent's life. Similarly, spouses did next best, but not as well as the family members, perhaps because they came to know the respondent later in life. Finally, friends did the worst, most likely because they probably have not known the respondent as long or as well as the other groups. Therefore, the notion that social distance affects personal knowledge has merit.

5. Recall of Word Associations

a. Recall of Word Associations by Respondents

The overall average number of correct matches by the respondents on all the word associations between Q1 and Q3 was 13.8 out of 20 (69%). The respondents fell anywhere in the continuum from 0 to 20 responses correct as shown in Figure 4. Of note is that 60 (58.3%) got 14 (70%) or more matches correct.

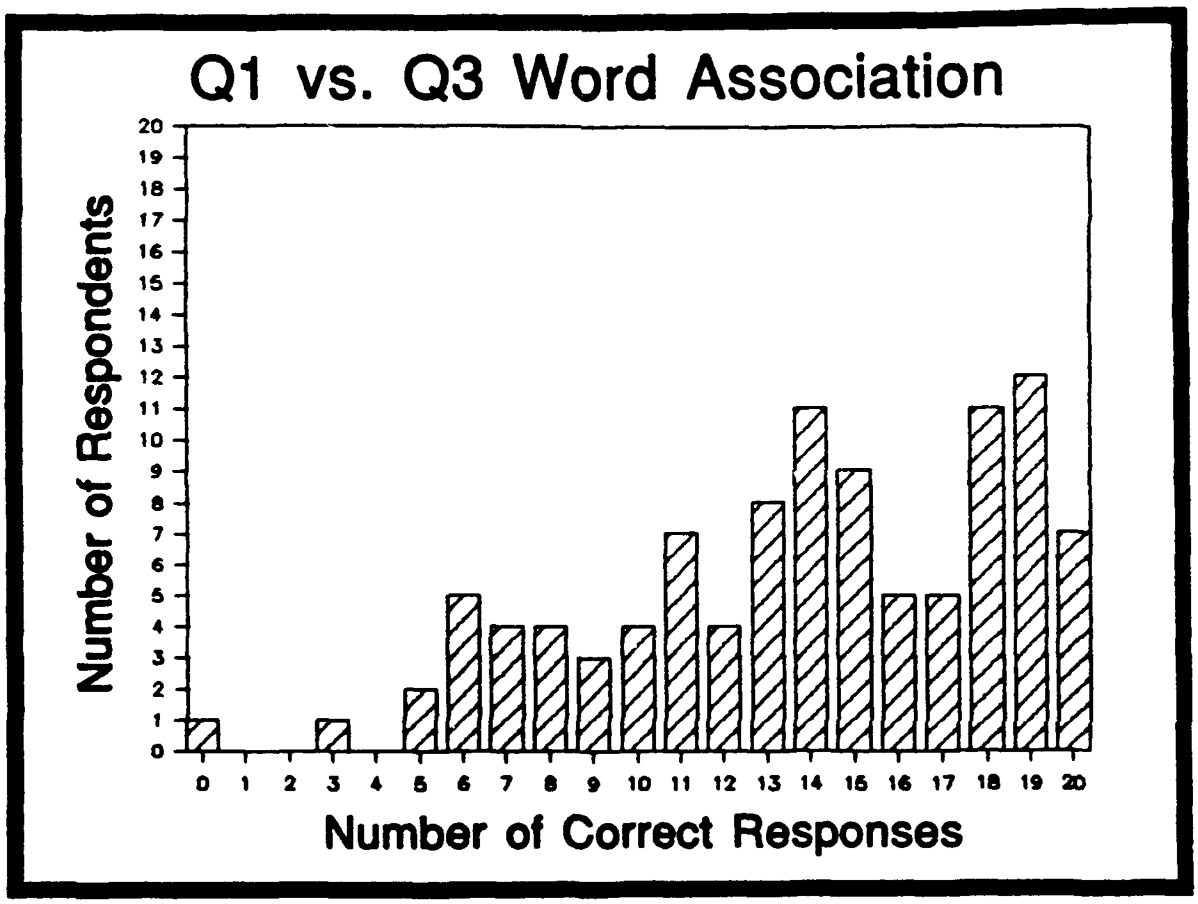

Figure 4. Q1 vs. Q3 Word Association

As expected, when first asked to generate both their cues and responses, the respondents on average were able to generate only 4.1 out of the 20 (20.5%). However, when presented with their list of cues they were able to generate responses albeit with some errors. Not one respondent requested to know what their theme was. Either there was no theme or the list of 20 cues made them remember their theme. While it was expected that few would need their theme to generate responses, it was surprising that not one, including the one who got none of his responses correct, requested his theme to help figure out the responses.

b. Matching of Word Associations by Significant-Others

The significant-others were first asked to guess the responses, without any help from the respondents, after having been given the cues. In this case, the significant-others were able to correctly match 5.1 out of 20 (25.5%). Then they were given the theme from the respondent. Only 39 of the 87 (44.8%) of the significant others used this information in an attempt to better their score. By using the theme, the significant-others improved their score by 3.9 correct responses on average. As a result, the significant-others overall were able to improve their score to 6.6 out of 20 (33%). Figure 5 shows the distribution of correct responses by the significant-others before being given the theme. Figure 6 shows the distribution of correct

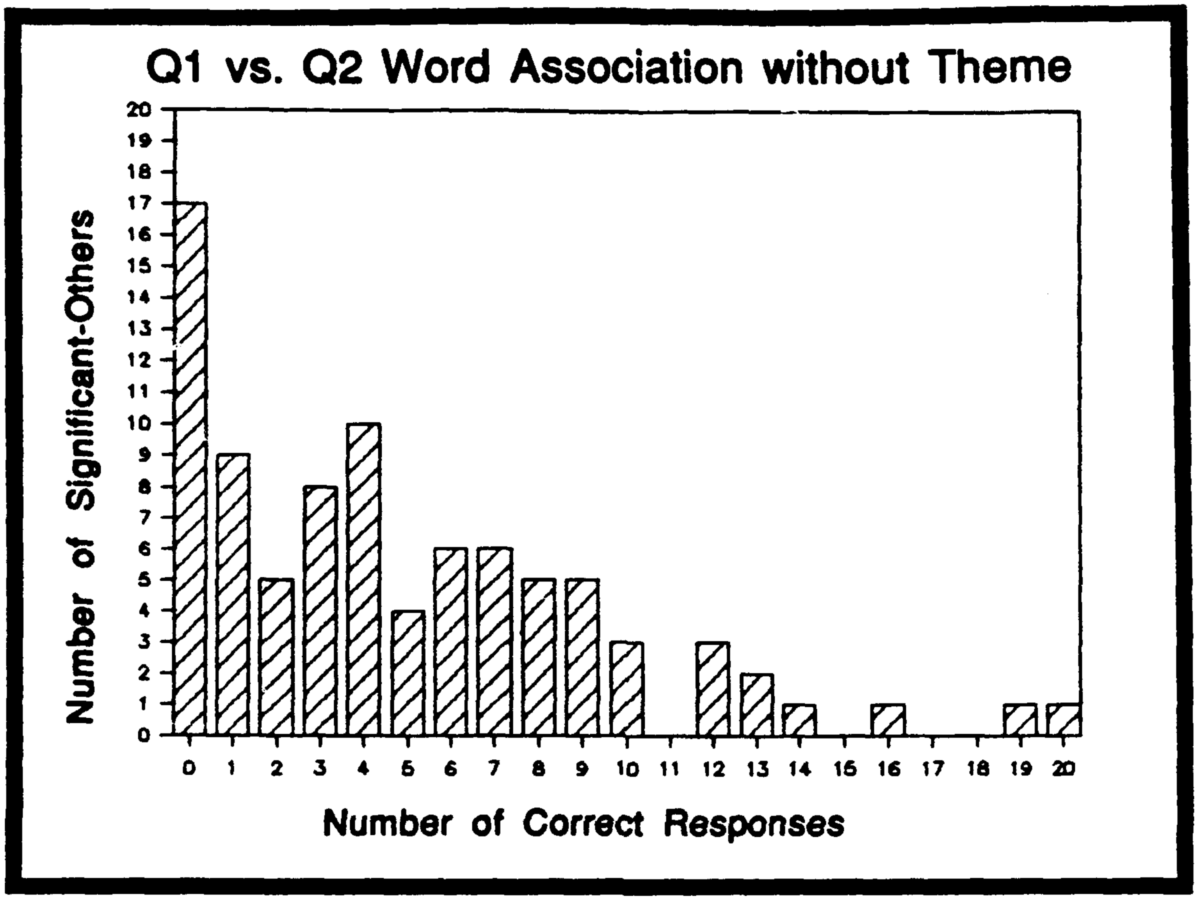

Figure 5. Q1 vs. Q2 Word Association without Theme

responses after being given the theme. It should be noted that only 39 significant-others attempted to use the theme to improve their score.

The distribution is skewed toward the lower end of the spectrum. However, there are a few outliers which are explained by the fact that the respondents chose trivial associations. For instance, the respondent of the

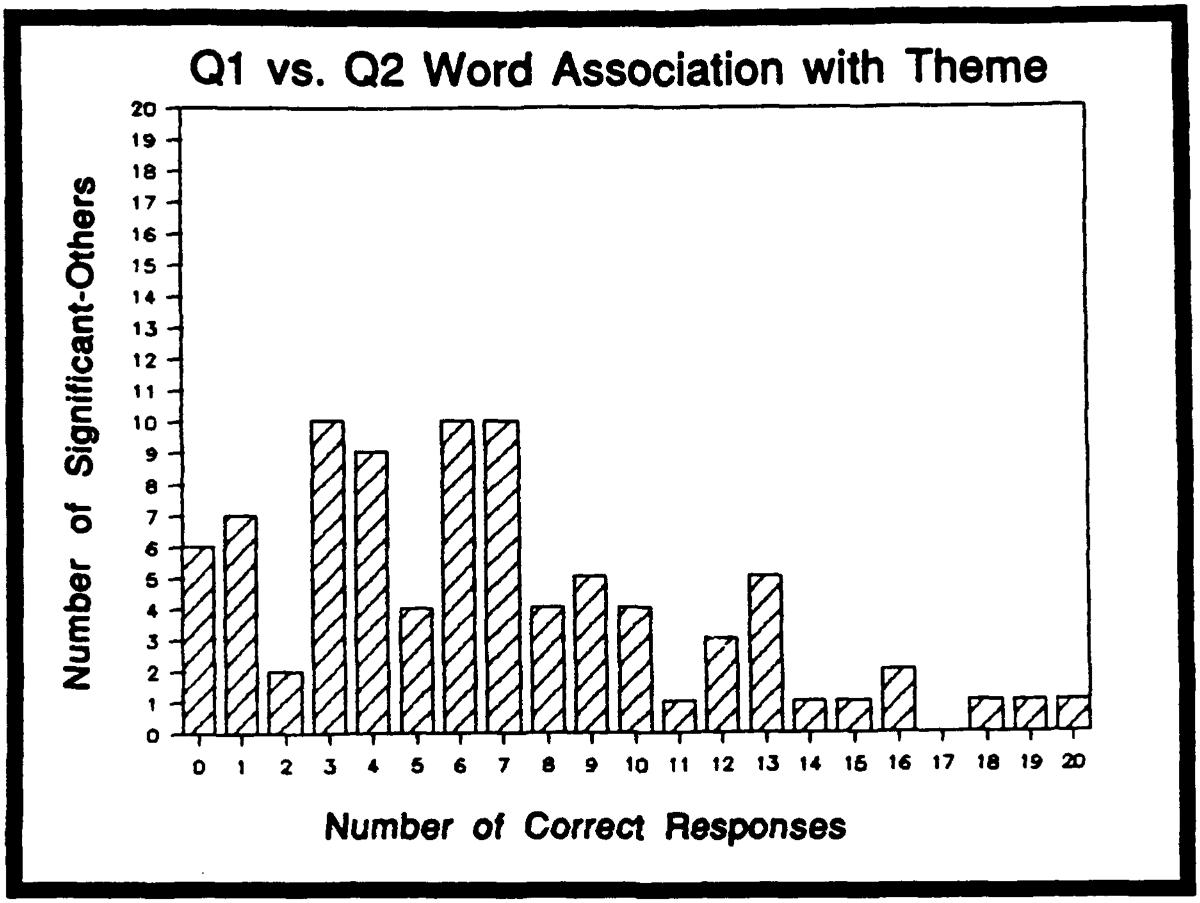

Figure 6. Q1 vs. Q2 Word Association with Theme

significant-other who got all 20 responses correct chose opposites as the theme.

Unlike the cognitive passwords, social closeness played no part in the ability of the significant-other to figure out correctly the responses. There was less than one correct response difference between the three groups--family members, spouses and friends.

6. Ranking of the Various Methods

The last task the respondents were asked to perform was to rank the various alternative methods of user authentication. First, they were asked to rank the five methods based on how easy each method was to remember. User-generated passwords were ranked first 50 times. Second was authentication by word association with 29 first place rankings. Third was cognitive passwords with 14 first place rankings. Passphrases were fourth with two first places. Finally, system-generated passwords (no distinction was made between random or pronounceable) were ranked last with no one choosing it as easiest to remember.

The respondents were then asked to rank the various methods according to how they liked them. The order was the same. User-generated passwords received 47 first place rankings, word association had 30, cognitive passwords had 16, passphrases had three, and one person liked system-generated passwords the best.

Eight respondents did not complete this part of the questionnaire. The rankings are summarized in Table 15.

TABLE 15

RANKINGS OF VARIOUS PASSWORD METHODS

RANKING BY EASE OF MEMORY:

METHOD	OVERALL RANK	AVERAGE SCORE
USER-GENERATED PASSWORDS	1	1.98
AUTHENTICATION BY WORD ASSOCIATION	2	2.41
COGNITIVE PASSWORDS	3	2.67
PASSPHRASES	4	3.45
SYSTEM-GENERATED PASSWORDS	5	4.46

RANKING BY HOW IT WAS LIKED:

METHOD	OVERALL RANK	AVERAGE SCORE
USER-GENERATED PASSWORDS	1	1.96
AUTHENTICATION BY WORD ASSOCIATION	2	2.39
COGNITIVE PASSWORDS	3	2.73
PASSPHRASES	4	3.38
SYSTEM-GENERATED PASSWORDS	5	4.54

VI. CONCLUSIONS AND RECOMMENDATIONS

A. DISCUSSION OF FINDINGS

1. Recall of Passwords and Passphrases

Over a three month period, only 27.2% of the respondents could recall the password that they had created themselves. As in previous research studies, this survey showed that as password length increased it became more difficult to remember (Table 3).

Similarly, only 12.7% of the respondents could remember their system-generated random alphanumeric password. However, it is important to note that of the respondents assigned a system-generated pronounceable password, 37.5% were able to recall it. These pronounceable passwords appear to be more secure than the user-generated passwords, perhaps because the pronounceable passwords are not related to the user's lifestyle. It was shown here, however, that even though unrelated, they are more memorable to the user than his self-generated password. It should be pointed out that not one respondent was able to remember the random alphanumeric password on his own. Among those who did recall it, 85.7%, had written it down.

21.4% of the 103 respondents were able to remember their passphrases. Most of the respondents, 77.7%, chose passphrases consisting of fewer than the minimum recommended

thirty characters (Porter, 1982). They still had little success in recalling the passphrase. Given that the accepted limit of short term human memory is seven characters, it is interesting to note that the percentage who remembered their passphrase was the same as those who remembered their user-generated password if it was eight characters in length: approximately 21%.

2. Recall of Cognitive Passwords

After three months, the respondents recalled an average of 74% of their cognitive passwords. Two of these respondents were able to recall all 20. When the fact-based cognitive passwords were analyzed separately, the recall averaged over 83%. The recall performance on the opinion-based cognitive passwords was somewhat lower than for the fact-based passwords. As a result, only 74.8% of the opinion-based cognitive passwords were recalled.

Recall of the cognitive passwords was noticeably better than for any of the previously described password alternatives. Overall, the findings support the notion that the ease of recall of cognitive passwords is superior to that of traditional passwords and slight modifications of that method.

The people who are socially close to the respondents (family members, spouses and friends), could guess no more than an average of 38% of the respondents' cognitive passwords. Only a few significant-others could legitimately

guess more than ten out of 20 responses due to previous exposure to a similar questionnaire. Two significant-others could not guess any of the 20 responses correctly.

When the guessing of fact-based cognitive passwords was analyzed separately from opinion-based ones, the results were as expected. People close to the respondent could guess fact-based cognitive passwords better than they could guess opinion-based ones. On average, the significant-others guessed 44.8% of the fact-based cognitive passwords while averaging only 32.5% for the opinion-based cognitive passwords.

The notion that people more socially close to the respondents are better guessers than those even slightly removed, was found to be true. The average number of correct guesses for family members was 12 (60%), while spouses were 8.2 (41%) and friends were 4.7 (23.5%).

3. Recall of Word Associations

After three months, the respondents recalled, on average, 69% of their word associations. Seven respondents remembered all 20 responses and almost a third remembered 90% or more of their responses. While there was success at the high end of the spectrum, there was a fairly uniform distribution of respondents remembering from 30% to 90%. An explanation for this distribution is that the respondents were given free reign in making up their word associations. Unlike the cognitive password section, in which all the

respondents answered the same questions, the word associations had various degrees of difficulty depending upon how challenging each respondent decided to make them.

Even with the wide variance, the average success rate was over twice that of the traditional user-generated password method. In comparison with the overall success rate of cognitive passwords, word associations were not as great (69% to 74%). However, there were almost twice as many respondents (30 to 17) scoring 90% or more correct responses on the word associations than on the cognitive passwords.

The significant-others, on average, could guess only 25.5% of the correct responses. Seventeen significant-others could not guess even one response correctly. There was a small percentage of significant-others (10.3%) who were able to guess correctly more than ten responses. As expected, when the respondents helped their significant-other by telling them what their theme was, the success rate improved. But only to 33%. There were still six significant-others who could not get any responses correct. Only 44.8% of the significant-others used information about the theme to improve their scores; not everyone did. Of the remaining 55.2% who did not use this information to improve their scores, the assumption cannot be made that they had figured out the theme since some of the respondents constructed word association lists without themes. This

shows that even though the significant-others saw all 20 cues at once--a luxury an intruder would not have--it was not obvious what the connection was among the cues.

Even with the theme, the significant-others failed to guess as many correct responses (33% to 38%) as they had in the cognitive passwords section. Also, unlike cognitive passwords, social closeness made no significant difference in the ability of the significant-others to figure out the responses.

4. <u>Ranking of the Various Methods</u>

When asked to rank the various methods as to how easy they were to remember, the respondents clearly chose user-generated passwords as the one that they thought was easiest. However, this method was one of the worst for recall by the respondents. Other than this, the rankings generally reflect how the respondents actually did in recalling their "passwords" from the different methods.

When the respondents ranked the methods by how they liked them, those that were user-oriented were ranked highest. Of interest is the fact that there was little difference between the two rankings. No method differed on the final score by even .1. This shows that the respondents may have interpreted that how they liked a certain method meant that it was easy to remember. This would explain why the respondents chose user-generated passwords as easiest to remember when in reality they were not.

5. Summary

These findings demonstrate that there are significant differences among the various password alternatives. Moreover, they show that both cognitive passwords and authentication by word association are methods that are easy for users to recall yet are difficult for others to guess, even by the people who know the users best.

B. COMPARISON TO OTHER STUDIES

1. Cognitive Passwords

Hulsey showed that cognitive passwords provided a better authentication method than traditional password systems. This study supports that conclusion. However, the survey group here did not provide the same clear-cut choice. Unlike Hulsey's survey group, there were outliers from both the respondents and the significant-others. Two explanations for the slight differences between these two studies are: (1) This survey group was larger, both in respondents (by 5.0%) and significant-others (by 11.5%). Even though these numbers are not large, neither were the differences in the studies. (2) Some of this survey group was the same as Hulsey's, so prior exposure to the cognitive password questions helped those significant-others do better this time around. Like Hulsey's study, this study showed that cognitive passwords were easy to remember and more difficult to figure out than user-generated passwords. However, this study shows that users still preferred the

traditional password method over cognitive passwords. Hulsey recommended that research be conducted into how authentication by word association compared to cognitive passwords. That research is discussed and compared with Smith's in the next section.

2. Word Association

Smith's research showed that after six months the four respondents in his survey group could recall 94% of their word associations (Smith, 1987). This is considerably higher than the 69% success rate after three months from this survey group. The difference in sizes of the two groups probably accounts for the difference in success rate. Smith had only two of his four respondents' significant-others try to guess the proper responses. They had a success rate of 45% and 50% respectively. He speculated that the lists from his other two respondents would be difficult to guess, unless some prior special knowledge about the respondent was known by the significant-other (Smith, 1987). The success rate of his significant-others was higher than the success rate of 25.5% (33% given the theme) by this survey group. Smith concluded that authentication by word association seemed promising for finding a better method for user authentication. The results of this study support his conclusion.

C. RECOMMENDATIONS

Several different user authentication methods were examined in this thesis. The desired security level of the organization and its access control policy need to be known before definitive advice can be given on choice of specific method. If an organization desires just to upgrade its traditional password system, without making radical changes, user-generated pronounceable passwords should be used. System-generated pronounceable passwords were proven easiest to remember. The one pitfall they had was that users dislike system-generated passwords; so by allowing the users to choose them, this password method should be more desirable to the user. Pronounceable passwords also offer a high degree of security as they are a mix of alphanumeric characters that do not form an actual word or phrase.

If an organization desires to change its present user authentication method to make it the best possible, authentication by word association should be chosen. A close second is cognitive passwords. Authentication by word association has been shown to be the most secure of the various methods discussed here. The 25.5% guess rate was lower than Hulsey's 27% guess rate for cognitive passwords. Also, users ranked it second to user-generated passwords as both easiest to remember and the one they liked. Even though the respondents, on average, did not respond to the word associations as well as they did on cognitive

passwords, almost twice as many respondents scored at the high end (90% or better) on the word associations than they did on cognitive passwords.

Both authentication by word association and cognitive passwords provide better security than traditional password systems. They are user-friendly and offer ease of memorability. Implementation of and continued research into these two methods should be encouraged.

APPENDIX

THESIS QUESTIONNAIRES

This appendix contains the three questionnaires Q1, Q2 and Q3.

THESIS SURVEY: PASSWORDS, PASSPHRASES AND AUTHENTICATION METHODS

PART A: PERSONAL INFORMATION

Please answer the following questions:

Sex (Circle one): Male Female

SMC No._______ or last 4 digits of your SSN________

Number of years of computer usage:_______

Type of computer(s) you have used prior to NPS (check any that apply):

Microcomputer______
Microcomputer linked to a mainframe_____
Mainframe terminal_____

PART B: PASSWORDS AND PASSPHRASES

For the purpose of this survey anytime you are requested to memorize something **do not write it down.** This is for all parts of this survey--passwords, passphrases, cognitive passwords and word association.

1. Please create and write in the boxes below a password, up to eight alphanumeric characters. Please memorize and safeguard it as you normally do your passwords. As with other parts of this survey, you will later be asked to recall what you have been requested to memorize.

|__|__|__|__|__|__|__|__|

2. How did you choose the password above? (Circle one)

A. A meaningful detail (name, date, number, etc.)
B. A combination of meaningful details (JIM1989, etc.)
C. A randomly chosen combination of characters
D. Other (please specify)_____________________________

3. The following password has been assigned to you for this study. Please memorize and safeguard it as you would any other password. This password is pronounceable, which may help you remember it. For instance, UN4TUNE8 would be unfortunate.

|__|__|__|__|__|__|__|__|

4. A passphrase is a string of up to 80 alphanumeric characters. Theoretically, it is more secure than a normal password since it is unlikely that someone will guess it. The passphrase can be silly like "Susie sells seashells by the seashore," or it can be a quotation or a common phrase. Please construct a passphrase of your choice in the space below. Please memorize it and safeguard it as you would any other password.

____

5. How did you choose your passphrase above? (Circle one)

   A. Nonsensical phrase that I can remember
   B. A quotation
   C. A piece of advice
   D. A common phrase
   E. Other (please specify)____________________________

PART C: COGNITIVE PASSWORDS

Cognitive passwords suggest the use of fact, interest and opinion-based cognitive data, that are known only to the user as an authentication mechanism. Please answer the following questions using a maximum of 20 characters.

1. What is the name of the elementary school from which you graduated? __________________

2. What is the first name of your favorite uncle? __________________

3. What is the first name of your best friend in high school? __________________

4. What is your mother's maiden name? __________________

5. What was the first name of your first boyfriend/ girlfriend? __________________

6. What was the name of your favorite class in high school? __________________

7. What is the name of your favorite music performer or group? __________________

9. What is the name of your favorite vacation place? __________________

10. If you could travel to any country in the world, which would it be? ________________

11. What is the last name of your favorite actor or actress? ________________

12. What is your favorite flower? ________________

13. What is your favorite dessert? ________________

14. What is your favorite vegetable? ________________

15. What is your favorite fruit? ________________

16. What is your favorite color? ________________

17. If you could change occupations, which new occupation would you choose? ________________

18. What is the name of your favorite restaurant ________________

19. What is the occupation of your father? ________________

20. What is the last name of your favorite college instructor? ________________

PART D: WORD ASSOCIATION

Another form of access control is a challenge-and-response query after a user has logged on. When the user correctly responds to the queries, the system ensures that it is the authorized user who has logged on. One such method is a series of word associations. Each user creates 20 word associations peculiar to him. For instance a user could decide to set up a table composed of queries that remind him of musical artists. The partial table is listed:

QUERY	RESPONSE
Virgin	Madonna
Deaf	Beethoven
Eliminator	ZZTop
Glasses	Elton_John
Lips	Mick_Jagger

So, after initial logon, the system would query: Glasses? The authentic user would then respond: Elton_John

1. Now construct a set of word associations for yourself. Please list 20 associations. While it is helpful for memory purposes to use one theme throughout it is not mandatory. Here are some other suggestions for possible themes: comic strips, authors, TV shows, movies, family members.

<u>QUERY</u>	<u>RESPONSE</u>
1. ____________	____________
2. ____________	____________
3. ____________	____________
4. ____________	____________
5. ____________	____________
6. ____________	____________
7. ____________	____________
8. ____________	____________
9. ____________	____________
10. ____________	____________
11. ____________	____________
12. ____________	____________
13. ____________	____________
14. ____________	____________
15. ____________	____________
16. ____________	____________
17. ____________	____________
18. ____________	____________
19. ____________	____________
20. ____________	____________

Theme (if any)________________________

2. Now rewrite the queries onto the survey for your spouse or friend to fill out. The instructions on how to administer the survey to your spouse or friend is on that survey, as well as the directions on where to return that survey when it is complete.

THESIS QUESTIONNAIRE--COGNITIVE PASSWORDS AND WORD ASSOCIATION

SMC NO.________ or last 4 digits of your SSN _________ ,
Relationship _______________ (to correlate with survey from class)

PART A: COGNITIVE PASSWORDS

Cognitive passwords suggest the use of fact, interest and opinion-based cognitive data, that are known only to the user as an authentication mechanism. Please answer the following questions, using a maximum of 20 characters, the way you think the person who gave you this survey answered them.

1. What is the name of the elementary school from which he/she graduated? _________________
2. What is the first name of his/her favorite uncle? _________________
3. What is the first name of his/her best friend in high school? _________________
4. What is his/her mother's maiden name? _________________
5. What was the first name of his/her first boyfriend/girlfriend? _________________
6. What was the name of his/her favorite class in high school? _________________
7. What is the name of his/her favorite music performer or group? _________________
8. What is his/her favorite type of music? _________________
9. What is the name of his/her favorite vacation place? _________________
10. If he/she could travel to any country in the world, which would it be? _________________
11. What is the last name of his/her favorite actor or actress? _________________
12. What is his/her favorite flower? _________________
13. What is his/her favorite dessert? _________________
14. What is his/her favorite vegetable? _________________

15. What is his/her favorite fruit? ________________

16. What is his/her favorite color? ________________

17. If he/she could change occupations, which new occupation would he/she choose? ________________

18. What is the name of his/her favorite restaurant? ________________

19. What is the occupation of his/her father? ________________

20. What is the last name of his/her favorite college instructor? ________________

PART B: WORD ASSOCIATION

A form of computer access control is a challenge-and-response query after a user has logged on. When the user correctly responds to the queries, the system ensures it is the authorized user who has logged on. One such method is the use of a series of word associations. Each user creates 20 word associations peculiar to him. For instance, a user could create queries that remind him of musical artists. For example:

QUERY	RESPONSE
Virgin	Madonna
Deaf	Beethoven
Eliminator	ZZTop
Glasses	Elton_John
Lips	Mick_Jagger

So, after initial logon, the system would query: Glasses?
The authentic user would respond: Elton_John

The person who presented you with this survey has created a table of 20 word associations. The queries are listed below and continue onto the next page. In column A of the responses, try to guess what the correct response to each query is supposed to be. When you are finished, ask if there was any theme to the 20 associations. Then he or she will tell you if their associations had a theme and if so what it was. Now try to see how many more correct responses you can get in column B.

	QUERY	COLUMN A RESPONSE	COLUMN B RESPONSE
1.			
2.			
3.			
4.			
5.			
6.			
7.			
8.			
9.			
10.			
11.			
12.			
13.			
14.			
15.			
16.			
17.			
18.			
19.			
20.			

Please return this form to either Mark Beedenbender (SMC No. 1749), your instructor or to Professor Zviran I-310 (or his mailbox on the second deck of Ingersoll).

THESIS SURVEY: PASSWORDS, PASSPHRASES AND AUTHENTICATION METHODS

PART A: PERSONAL INFORMATION

SMC No._______ or last 4 digits of your SSN_______

PART B: PASSWORDS AND PASSPHRASES

Earlier this quarter you were asked to memorize several passwords and a passphrase. This survey will test your recall of those passwords and passphrase.

1. Please create and write in the boxes below a password, up to eight alphanumeric characters. Use the same password you used at the beginning of the quarter.

   |__|__|__|__|__|__|__|__|

2. How did you remember the password above? (Circle one)

   A. I did write it down because I knew that would be the only way I could remember it.
   B. I just remembered it.
   C. It is the only password I ever use so it was easy to remember.
   D. Other (please specify) ____________________________

3. Please enter the password that was assigned to you for this study in the boxes below.

   |__|__|__|__|__|__|__|__|

4. How did you remember the password above? (Circle one)

   A. Since it was pronounceable, it was easy to remember.
   B. I just remembered it.
   C. I did write it down because I knew that would be the only way I could remember it.
   D. Other (please specify) ____________________________

5. Please enter the passphrase that you chose at the beginning of the quarter in the space below.

   ____

6. How did you remember your passphrase above? (Circle one)

   A. I did write it down because I knew that would be the only way I could remember it.
   B. I just remembered it.
   C. It's a phrase I use over and over again so it was easy to remember.
   D. Other (please specify) ____________________

PART C: COGNITIVE PASSWORDS

Please answer the following questions using a maximum of 20 characters.

1. What is the name of the elementary school from which you graduated? ____________
2. What is the first name of your favorite uncle? ____________
3. What is the first name of your best friend in high school? ____________
4. What is your mother's maiden name? ____________
5. What was the first name of your first boyfriend/ girlfriend? ____________
6. What was the name of your favorite class in high school? ____________
7. What is the name of your favorite music performer or group? ____________
8. What is your favorite type of music? ____________
9. What is the name of your favorite vacation place? ____________
10. If you could travel to any country in the world, which would it be? ____________
11. What is the last name of your favorite actor or actress? ____________
12. What is your favorite flower? ____________
13. What is your favorite dessert? ____________
14. What is your favorite vegetable? ____________

15. What is your favorite fruit? ____________

16. What is your favorite color? ____________

17. If you could change occupations, which new occupation would you choose? ____________

18. What is the name of your favorite restaurant? ____________

19. What is the occupation of your father? ____________

20. What is the last name of your favorite college instructor? ____________

PART D: WORD ASSOCIATION

1. Try to reconstruct the set of word associations that you made for yourself at the beginning of the quarter. There were 20 associations.

	QUERY	RESPONSE
1.	______	______
2.	______	______
3.	______	______
4.	______	______
5.	______	______
6.	______	______
7.	______	______
8.	______	______
9.	______	______
10.	______	______
11.	______	______
12.	______	______
13.	______	______
14.	______	______
15.	______	______
16.	______	______
17.	______	______
18.	______	______

19.____________ __________________

20.____________ __________________

THESIS QUESTIONNAIRE--AUTHENTICATION BY WORD ASSOCIATION

SMC NO.__________ or your last 4 digits of your SSN _________

Listed below are the 20 queries that you created at the beginning of the quarter. In column A of the responses, try to guess what the correct response to each query is supposed to be. If you are unable to respond to all of the queries ask the person who is administering the questionnaire if there was any theme for the 20 associations. Now try to see how many more correct responses you can get writing your responses in column B.

QUERY	COLUMN A RESPONSE	COLUMN B RESPONSE
1. ________	________	________
2. ________	________	________
3. ________	________	________
4. ________	________	________
5. ________	________	________
6. ________	________	________
7. ________	________	________
8. ________	________	________
9. ________	________	________
10. ________	________	________
11. ________	________	________
12. ________	________	________
13. ________	________	________
14. ________	________	________
15. ________	________	________
16. ________	________	________
17. ________	________	________
18. ________	________	________
19. ________	________	________
20. ________	________	________

1. In this survey, you saw five different authentication methods. Please rank these methods according to how easy it was to remember the method. Use a ranking scale where "1" is the easiest to remember, while "5" would be the most difficult to remember. Use each number, 1 through 5, only once.

	RANK
Personally selected passwords	________
System generated passwords	________
Passphrases	________
Cognitive passwords	________
Authentication by word association	________

2. Now rank the methods according to how you liked them. This time "1" would stand for your most favorite, while your least favorite would be "5." Again, use each number, 1 through 5, only once.

	RANK
Personally selected passwords	________
System generated passwords	________
Passphrases	________
Cognitive passwords	________
Authentication by word association	________

LIST OF REFERENCES

Ahituv, N., Lapid, Y. and Neumann, S. (1987), "Verifying the Authentication of an Information System User," Computers and Security, Vol. 6, No. 2, pp. 152-157.

Barton, B.F. and Barton, M.S. (1984), "User-Friendly Password Methods for Computer-Mediated Information Systems," Computers and Security, Vol. 3, No. 3, pp. 186-195.

Denning, D.E. and Denning, P.J. (1979), "Security," Computing Surveys, Vol. 11, No. 3 (September), pp. 227-247.

Evans, A., Kantrowitz, W. and Weiss, E. (1974), "A User Authentication Scheme Not Requiring Secrecy in the Computer," Communications of the ACM, Vol. 17, No. 8, pp. 437-442.

Haga, W.J., Hulsey, J.D. and Zviran, M. (1989), "Cognitive Passwords: From Theory to Practice," Working Paper No. 89-06, pp. 1-16, Naval Postgraduate School, Monterey, California.

Hoffer, J.A. and Straub, D.W. (1989), "The 9 to 5 Underground: Are You Policing Computer Crimes?", Sloan Management Review, Summer, pp. 35-43.

Landwehr, C.E. (1981), "Formal Models of Computer Security," Computing Surveys, Vol. 13, No. 3 (September), pp. 247-278.

Menkus, B. (1980), "Understanding the Use of Passwords," Computers and Security, Vol. 7, No. 2, (April), pp. 132-136.

Miller, G.A. (1956), "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," The Psychological Review, Vol. 63, (March), pp. 81-97.

Morris, R. and Thompson, K. (1979), "Password Security: A Case History," Communications of the ACM, Vol. 22, No. 11, pp. 594-597.

Panns, R. and Herschberg, I.S. (1987), "Computer Security: The Long Road Ahead," Computers and Security, Vol. No. 5, pp. 403-416.

Parker, D.B. and Nycum, S.H. (1984), "Computer Crime," Communications of the ACM, Vo. 27, No. 4 (April), pp. 313-321.

Porter, S.N. (1982), "A Password Extension for Improved Human Factors," Computers and Security, Vol. 1, No. 1, pp. 54-56.

Pfleeger, C.P. (1989), Security in Computing, pp. 75-83, Prentice-Hall, Inc.

Smith, S.L. (1987), "Authenticating Users by Word Association," Computers and Security, Vol. 6, No. 6, pp. 464-470.

Spender, J.C. (1987), "Identifying Computer Users with Authentication Devices (Tokens)," Computers and Security, Vol 6., No. 6, pp. 385-395.

Ware, W.H. (1984), "Information Systems Security and Privacy," Communications of the ACM, Vol. 27, No. 4 pp. 315-321.

Wood, C.C. (1983), "Effective Information System Security with Password Controls," Computers and Security, Vol. 2, No. 1, pp. 5-10.

Wood, C.C. (1987), "The Human Immune System as an Information System Security Reference Model," Computers and Security, Vol. 6, pp. 511-516.

<u>INITIAL DISTRIBUTION LIST</u>

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22304-6145	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93943-5002	2
3.	Department Chairman, Code AS Department of Administrative Sciences Naval Postgraduate School Monterey, California 93943-5000	1
4.	Prof. William J. Haga, Code AS/Hg Department of Administrative Sciences Naval Postgraduate School Monterey, California 93943-5000	1
5.	Prof. Moshe Zviran, Code AS/Zv Department of Administrative Sciences Naval Postgraduate School Monterey, California 93943-5000	1
6.	Computer Technology Curricular Office, Code 37 Naval Postgraduate School Monterey, California 93943-5000	1
7.	Director, Computer Center, Code 0141 Naval Postgraduate School Monterey, California 93943-5000	1
8.	Dr. Carol Taylor 990 Benito Court Pacific Grove, California 93950	1
9.	LT John D. Hulsey, U.S. Navy 522 Corey Lane Middletown, Rhode Island 02840	1
10.	LT Mark G. Beedenbender, U.S. Navy 1625 Dogwood Lane Lynchburg, Virginia 24503	2